

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/723,365B
Source: FFWO
Date Processed by STIC: 1-30-05

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:55

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

```

3 <110> APPLICANT: VAN DEN BOOM, DIRK
4   BOCKER, SEBASTIAN
6 <120> TITLE OF INVENTION: FRAGMENTATION-BASED METHODS AND SYSTEMS FOR SEQUENCE
7   VARIATION DETECTION AND DISCOVERY
9 <130> FILE REFERENCE: SEQ-2073-UT
11 <140> CURRENT APPLICATION NUMBER: 10/723,365B
12 <141> CURRENT FILING DATE: 2003-11-26
14 <150> PRIOR APPLICATION NUMBER: 60/429,895
15 <151> PRIOR FILING DATE: 2002-11-27
17 <160> NUMBER OF SEQ ID NOS: 85
19 <170> SOFTWARE: PatentIn Ver. 3.2
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 7
23 <212> TYPE: PRT
24 <213> ORGANISM: Artificial Sequence
26 <220> FEATURE:
27 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
28   peptide
30 <400> SEQUENCE: 1
31 Pro Phe His Leu Leu Val Tyr
32   1           5
35 <210> SEQ ID NO: 2
36 <211> LENGTH: 5
37 <212> TYPE: PRT
38 <213> ORGANISM: Artificial Sequence
40 <220> FEATURE:
41 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
42   peptide
44 <220> FEATURE:
45 <221> NAME/KEY: MOD_RES
46 <222> LOCATION: (5)
47 <223> OTHER INFORMATION: Any amino acid except pro or arg
49 <400> SEQUENCE: 2
W--> 50 Ile Glu Gly Arg Xaa
51   1           5
54 <210> SEQ ID NO: 3
55 <211> LENGTH: 5
56 <212> TYPE: PRT
57 <213> ORGANISM: Artificial Sequence
59 <220> FEATURE:
60 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
61   peptide
63 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:55

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

```

64 <221> NAME/KEY: MOD_RES
65 <222> LOCATION: (5)
66 <223> OTHER INFORMATION: Any amino acid except pro or arg
68 <400> SEQUENCE: 3
W--> 69 Ile Asp Gly Arg Xaa
70   1           5
73 <210> SEQ ID NO: 4
74 <211> LENGTH: 5
75 <212> TYPE: PRT
76 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
80   peptide
82 <220> FEATURE:
83 <221> NAME/KEY: MOD_RES
84 <222> LOCATION: (5)
85 <223> OTHER INFORMATION: Any amino acid except pro or arg
87 <400> SEQUENCE: 4
W--> 88 Ala Glu Gly Arg Xaa
89   1           5
92 <210> SEQ ID NO: 5
93 <211> LENGTH: 5
94 <212> TYPE: PRT
95 <213> ORGANISM: Artificial Sequence
97 <220> FEATURE:
98 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
99   peptide
101 <220> FEATURE:
102 <221> NAME/KEY: MOD_RES
103 <222> LOCATION: (2)
104 <223> OTHER INFORMATION: Variable amino acid
106 <220> FEATURE:
107 <221> NAME/KEY: MOD_RES
108 <222> LOCATION: (5)
109 <223> OTHER INFORMATION: Variable amino acid
111 <400> SEQUENCE: 5
W--> 112 Pro Xaa Gly Pro Xaa
113   1           5
116 <210> SEQ ID NO: 6
117 <211> LENGTH: 49
118 <212> TYPE: DNA
119 <213> ORGANISM: Artificial Sequence
121 <220> FEATURE:
122 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
123   primer
125 <400> SEQUENCE: 6
126 cagtaatacg actcactata gggagaaggc tccccagcaa gacggactt      49
129 <210> SEQ ID NO: 7
130 <211> LENGTH: 28

```

RAW SEQUENCE LISTING

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:55

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

```

131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
136     primer
138 <400> SEQUENCE: 7
139 aggaagagag cgcctcggca aagtacac                28
142 <210> SEQ ID NO: 8
143 <211> LENGTH: 340
144 <212> TYPE: DNA
145 <213> ORGANISM: Artificial Sequence
147 <220> FEATURE:
148 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
149     amplicon sequence
151 <400> SEQUENCE: 8
152 gggagaaggc tccccagcaa gacggacttc ttcaaaaaca tcatgaactt catagacatt 60
153 gtggccatca ttcttattt catcacgctg ggcaccgaga tagctgagca ggaaggaaac 120
154 cagaagggcg agcaggccac ctccctggcc atcctcaggg tcatccgctt ggtaagggtt 180
155 tttagaatct tcaagctctc ccgccactct aagggcctcc agatcctggg ccagaccctc 240
156 aaagctagta tgagagagct agggctgctc atctttttcc tcttcacggtt gggtcatcctg 300
157 ttttctagtg cagtgtactt tgccgaggcg ctctcttctc          340
160 <210> SEQ ID NO: 9
161 <211> LENGTH: 23
162 <212> TYPE: DNA
163 <213> ORGANISM: Artificial Sequence
165 <220> FEATURE:
166 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
167     primer
169 <400> SEQUENCE: 9
170 cccagtcacg acgttgtaaa acg                23
173 <210> SEQ ID NO: 10
174 <211> LENGTH: 23
175 <212> TYPE: DNA
176 <213> ORGANISM: Artificial Sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
180     primer
182 <400> SEQUENCE: 10
183 agcggataac aatttcacac agg                23
186 <210> SEQ ID NO: 11
187 <211> LENGTH: 117
188 <212> TYPE: DNA
189 <213> ORGANISM: Artificial Sequence
191 <220> FEATURE:
192 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
193     amplicon sequence
195 <400> SEQUENCE: 11
196 cccagtcacg acgttgtaaa acgtccaggg aggactcacc atgggcattt gattgcagag 60
197 cagctccgag tccatccaga gcttcctgca gtcacctgtg tgaaattgtt atccgct    117

```

RAW SEQUENCE LISTING

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:55

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

200 <210> SEQ ID NO: 12
 201 <211> LENGTH: 21
 202 <212> TYPE: DNA
 203 <213> ORGANISM: Artificial Sequence
 205 <220> FEATURE:
 206 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
 207 nucleotide sequence
 209 <220> FEATURE:
 210 <221> NAME/KEY: modified_base
 211 <222> LOCATION: (1)..(3)
 212 <223> OTHER INFORMATION: a, c, g, t, unknown, or other
 214 <220> FEATURE:
 215 <221> NAME/KEY: modified_base
 216 <222> LOCATION: (8)..(10)
 217 <223> OTHER INFORMATION: a, c, g, t, unknown, or other
 219 <220> FEATURE:
 220 <221> NAME/KEY: modified_base
 221 <222> LOCATION: (12)..(14)
 222 <223> OTHER INFORMATION: a, c, g, t, unknown, or other
 224 <220> FEATURE:
 225 <221> NAME/KEY: modified_base
 226 <222> LOCATION: (19)..(21)
 227 <223> OTHER INFORMATION: a, c, g, t, unknown, or other
 229 <400> SEQUENCE: 12

W--> 230 nnnactgnnn mnnntgacnn n

21

233 <210> SEQ ID NO: 13
 234 <211> LENGTH: 583
 235 <212> TYPE: DNA
 236 <213> ORGANISM: Artificial Sequence
 238 <220> FEATURE:
 239 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic amplicon
 240 sequence
 242 <400> SEQUENCE: 13
 243 cttcagtgtct cacaccgacc ctatgagtgg ggcggtcaaa ctgtcccat tttacacaca 60
 244 gggaaactta gtgaatggca aggctgggtt tgagcccagc tctattgcc ccaaagataa 120
 245 ggctccattc cctgctccat ttcccaggca taggggacttg tagggggctg gaaccccagg 180
 246 atcaactctg ggctcagagg gccccagcaa taagtgtactg ttgattactc ctgatcccaa 240
 247 agctgacttc aggcaagctc cttggagggtc gcagcccctt cttgctatgc ccagtggcaa 300
 248 tgatgttcat aatcccactc ctcagtgcag ggttccacta agaaccatg atctcctacc 360
 249 tcaaattggac ctcatgcttt ctgagtaagc ctccctcagc tttctgggtc cctcactccc 420
 250 cccacccact gcaatgactt cttcaggcct tccctgccat cctcaaactc ccagctgccc 480
 251 cctcctgtct accttccact tccctctcca cacacaacct gcttaccaga gagctgagca 540
 252 gagccaccaa cagaacttcc ccccccacgtc gctgctccca gtc 583
 255 <210> SEQ ID NO: 14
 256 <211> LENGTH: 483
 257 <212> TYPE: DNA
 258 <213> ORGANISM: Mycobacterium abscessus
 260 <400> SEQUENCE: 14
 261 acgggtgagt aacacgtggg tgatctgccc tgcactctgg gataagcctg ggaaactggg 60

RAW SEQUENCE LISTING

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:55

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

```

262 tctaataaccg gataggacca cacacttcat ggtgagtggg gcaaagcttt tgcggtgtgg 120
263 gatgagcccg cggcctatca gcttgttggg ggggtaatgg cccaccaagg cgacgacggg 180
264 tagccggcct gagaggggtga ccggccacac tgggactgag atacggccca gactcctacg 240
265 ggaggcagca gtggggaata ttgcacaatg ggcgcaagcc tgatgcagcg acgccgcgtg 300
266 agggatgacg gccttcgggt tgtaaacctc ttccagtagg gacgaagcga aagtgacggt 360
267 acctacagaa gaaggaccgg ccaactacgt gccagcagcc gcggtataac gtaggggtccg 420
268 agcgttgtcc ggaattactg ggcgtaaaga gctcgtaggg ggtttgtcgc gttgttcgtg 480
269 aaa 483
272 <210> SEQ ID NO: 15
273 <211> LENGTH: 495
274 <212> TYPE: DNA
275 <213> ORGANISM: Mycobacterium avium
277 <400> SEQUENCE: 15
278 acgggtgagt aacacgtggg caatctgccc tgcacttcgg gataagcctg ggaaactggg 60
279 tctaataaccg gataggacct caagacgcat gtcttctggg ggaaagcttt tgcggtgtgg 120
280 gatgggccccg cggcctatca gcttgttggg ggggtgacgg cctaccaagg cgacgacggg 180
281 tagccggcct gagaggggtg ccggccacac tgggactgag atacggccca gactcctacg 240
282 ggaggcagca gtggggaata ttgcacaatg ggcgcaagcc tgatgcagcg acgccgcgtg 300
283 ggggatgacg gccttcgggt tgtaaacctc ttccaccatc gacgaagggtc cgggttttct 360
284 cggattgacg gtaggtggag aagaagcacc ggccaactac gtgccagcag ccgcggtaat 420
285 acgtagggtg cgagcgttgt ccggaattac tgggcgtaaa gagctcgtag gtggtttgtc 480
286 gcgttgttcg tgaaa 495
289 <210> SEQ ID NO: 16
290 <211> LENGTH: 495
291 <212> TYPE: DNA
292 <213> ORGANISM: Mycobacterium celatum
294 <400> SEQUENCE: 16
295 acgggtgagt aacacgtggg tgatctgccc tgcacttcgg gataagcctg ggaaactggg 60
296 tctaataaccg gataggacca tgggatgcat gtcttgtggg ggaaagcttt tgcggtgtgg 120
297 gatgggccccg cggcctatca gcttgttggg ggggtgatgg cctaccaagg cgacgacggg 180
298 tagccggcct gagaggggtg ccggccacac tgggactgag atacggccca gactcctacg 240
299 ggaggcagca gtggggaata ttgcacaatg ggcgcaagcc tgatgcagcg acgccgcgtg 300
300 ggggatgacg gccttcgggt tgtaaacctc ttccaccatc gacgaagctg ccggttttcc 360
301 ggtggtgacg gtaggtggag aagaagcacc ggccaactac gtgccagcag ccgcggtaat 420
302 acgtagggtg cgagcgttgt ccggaattac tgggcgtaaa gagctcgtag gtggtttgtc 480
303 gcgttgttcg tgaaa 495
306 <210> SEQ ID NO: 17
307 <211> LENGTH: 483
308 <212> TYPE: DNA
309 <213> ORGANISM: Mycobacterium fortuitum
311 <400> SEQUENCE: 17
312 acgggtgagt aacacgtggg tgatctgccc tgcactttgg gataagcctg ggaaactggg 60
313 tctaataaccg aatatgacca cgcgttcat ggtgtgtggg ggaaagcttt tgcggtgtgg 120
314 gatgggccccg cggcctatca gcttgttggg ggggtaatgg cctaccaagg cgacgacggg 180
315 tagccggcct gagaggggtga ccggccacac tgggactgag atacggccca gactcctacg 240
316 ggaggcagca gtggggaata ttgcacaatg ggcgcaagcc tgatgcagcg acgccgcgtg 300
317 agggatgacg gccttcgggt tgtaaacctc ttccaatagg gacgaagcgc aagtgacggt 360
318 acctatagaa gaaggaccgg ccaactacgt gccagcagcc gcggtataac gtaggggtccg 420
319 agcgttgtcc ggaattactg ggcgtaaaga gctcgtaggg ggtttgtcgc gttgttcgtg 480

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/723,365B

DATE: 01/30/2005
TIME: 12:03:56

Input Set : A:\Seq2073u.app
Output Set: N:\CRF4\01302005\J723365B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 5

Seq#:3; Xaa Pos. 5

Seq#:4; Xaa Pos. 5

Seq#:5; Xaa Pos. 2,5

Seq#:12; N Pos. 1,2,3,8,9,10,12,13,14,18,20,21

Seq#:32; N Pos. 821

VERIFICATION SUMMARY

DATE: 01/30/2005

PATENT APPLICATION: US/10/723,365B

TIME: 12:03:56

Input Set : A:\Seq2073u.app

Output Set: N:\CRF4\01302005\J723365B.raw

L:50 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:69 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:88 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:112 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:230 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:596 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:780